Assessment of Cartilage Repair Quality With the International Cartilage Repair Society Score and the Oswestry Arthroscopy Score

Paatela, Teemu

2020-03

Paatela , T , Vasara , A , Nurmi , H , Kautiainen , H & Kiviranta , I 2020 , ' Assessment of Cartilage Repair Quality With the International Cartilage Repair Society Score and the Oswestry Arthroscopy Score ' , Journal of Orthopaedic Research , vol. 38 , no. 3 , pp. 555-562 . https://doi.org/10.1002/jor.24490

http://hdl.handle.net/10138/328472
https://doi.org/10.1002/jor.24490

acceptedVersion

Teemu Paatela ORCID iD: 0000-0002-6389-051X

Assessment of Cartilage Repair Quality with The International Cartilage Repair Society Score and The Oswestry Arthroscopy Score

Authors:

Teemu Paatela, MD (Corresponding author), Helsinki University Hospital, PL 900, 00029 HUS, Helsinki, Finland, +358-40-5864429, Teemu.Paatela@helsinki.fi

Anna Vasara, MD, PhD, Helsinki University Hospital, PL 900, 00029 HUS, Helsinki, Finland

Heikki Nurmi, MD, Jyväskylä Central Hospital, Keskussairaalantie 19, 40620 Jyväskylä, Finland

Hannu Kautiainen, MSc, Primary Health Care Unit, Kuopio University Hospital, Finland, and

Folkhälsan Research Center, Helsinki, Finland

Ilkka Kiviranta, MD, PhD, Helsinki University Hospital, PL 900, 00029 HUS, Helsinki, Finland, and University of Helsinki, PL 266, 00029 HUS, Helsinki, Finland

Contributions

Conception and design of the study: Teemu Paatela, Anna Vasara, Heikki Nurmi, Hannu Kautiainen, Ilkka Kiviranta. Data acquisition: Teemu Paatela, Anna Vasara, Heikki

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jor.24490.

Nurmi. Analysis and interpretation of data: Teemu Paatela, Anna Vasara, Hannu Kautiainen, Ilkka Kiviranta. Provision of study patients: Heikki Nurmi, Ilkka Kiviranta. Statistical expertise: Hannu Kautiainen. All authors participated in drafting and critically revising the article and final approval. Teemu Paatela takes responsibility for the integrity of the work as a whole, from inception to finished article.

Running Title (max Running title of 5 words or less.):

Arthroscopic scoring of cartilage repair

Abstract

The International Cartilage Repair Society (ICRS) score and the Oswestry Arthroscopic Score (OAS) have been validated to evaluate repair tissue quality. However, the performance of these scores have not been studied in typical patients undergoing cartilage repair and who have lesions of varying size. In this study, we compared the performance of the ICRS and the OAS scores and analyzed the effect of lesion characteristics on the performance of these two scores. Cartilage repair quality was assessed in a total of 104 arthroscopic observations of cartilage repair sites of the knee in 62 patients after autologous chondrocyte implantation (ACI). Two observers scored the repair areas independently with the ICRS and the OAS scores. The performance of both scores was evaluated according to internal consistency and inter-rater reliability and correlation between the scores. The frequency and proportion of disagreements were analyzed according to the repair site area and the given score. The correlation between the scores was good (r=0.91, 95% CI: 0.87-0.94). Both scores showed moderate internal consistency and inter-rater reliability. Cronbach's alpha was 0.88 (95% CI: 0.80-0.92) for the ICRS score and 0.79 (95% CI: 0.70-0.86) for the OAS score. The intraclass

correlation coefficient was 0.89 (95% CI: 0.84-0.92) for the ICRS and 0.81 (95% CI: 0.74-0.87) for the OAS scores. The frequency and proportion of disagreements were higher in larger repair sites. In arthroscopic use, both ICRS and OAS scores perform similarly, however, their reliability deteriorates as the lesion size increases.

Key words (3-6 keywords)

International Cartilage Repair Society score, Oswestry Arthroscopic Score, Arthroscopy, Cartilage repair, Repair tissue quality, Autologous chondrocyte implantation

Introduction

Articular cartilage lesions are a potential cause of knee pain and functional impairment. Cartilage repair aims to fill the defect, restore the properties of the original cartilage, and thereby reduce symptoms.[1; 2] Higher quality of repair tissue seems to lead to better clinical outcomes.[3; 4] Therefore, different cartilage repair techniques are evaluated according to the quality of the repair tissue which they are capable to produce.[5] Cartilage surfaces can be visualized with an arthroscope, and probing of the surfaces may help distinguish between pathological and normal cartilage.[6; 7] Therefore, arthroscopy can be used to evaluate the outcome of cartilage repair.[1; 2] The International Cartilage Repair Society (ICRS) score and the Oswestry Arthroscopy Score (OAS) were designed to evaluate arthroscopically the quality of cartilage repair.[5; 8] These scores evaluate the quality based on the following repair tissue properties: lesion fill, integration to the surrounding cartilage, appearance, and feeling on arthroscopic probing. These scores are used to provide important interim results before the clinical outcome can be assessed.[8] Additionally, macroscopic assessment of repair quality could be the primary outcome in trials in which a reliable clinical outcome is unavailable, such as in preclinical and

feasibility trials.[8] Both scores have been widely used in various studies evaluating cartilage repair outcome.[1; 9-12]

Both the ICRS and OAS scores have been validated for assessing repair quality arthroscopically after autologous chondrocyte implantation (ACI).[8; 10] However, as emphasized by many authors, the measures of validity are not necessarily generalizable. They are valid only in similar material in which the score performance was measured.[13-16] Importantly, reliability of a test is not an unchanged constant. Instead, the properties of the sample (in this case the cartilage lesions) may affect the reliability measures of a test.[15-17] Thus, the properties of cartilage lesions might affect the reliability of the ICRS and the OAS scores. Therefore, previous studies actually validate the use of ICRS and OAS scores for assessing specific types of lesions, i.e., lesions with an area of approximately 2.5 cm^2 and repaired with ACI.[10; 18] Furthermore, patients with such lesions are not considered to represent typical patients undergoing cartilage repair surgery of the knee, since patients typically have a larger lesion area.[19] Thus, the reliability of both scores to evaluate repair tissue quality in larger cartilage lesions is unknown. Furthermore, previous validation studies have shown limitations in score performance, but the reasons for these deficiencies remain unclear.[8; 10; 20] Because the ICRS and OAS scores are used to assess outcome of cartilage repair in various types of cartilage lesions, it is important to understand how the lesion characteristics might affect the reliability of both scores and what reasons might explain the previously observed limitations in score performance.

We designed this study to evaluate the reliability of the scores in typical patients undergoing cartilage repair and who have lesions of varying size. We compared the

performance of the ICRS and the OAS scores and analyzed the effect of lesion characteristics on the performance of these scores.

Patients and Methods

Diagnostic study, Level of Evidence III

Study Design and Study Subjects

In this diagnostic study, we evaluated cartilage repair tissue quality of 62 patients returning for second-look arthroscopies after cartilage repair of the knee with ACI. The repair procedures were performed between 1997 and 2008 in Jyväskylä Central Hospital, Finland. The arthroscopic evaluations were performed 0.3 to 11.4 years after the cartilage repair. The procedures were visually documented with still images at the time of the second-look arthroscopies. The documentation of these arthroscopies was arranged in randomly named files. The images and the operation notes of the second-look arthroscopies were used to assess the repair sites according to both ICRS (protocol A) and OAS scores (Table 1). Two orthopedic surgeons (TP and AV), familiar with cartilage repair techniques and who routinely perform knee arthroscopies, independently graded the lesions from the available images. The stiffness of the repair tissue was evaluated according to the description of the operation notes and the images. The two orthopedic surgeons grading the lesions were not involved in the repair surgeries or the second-look arthroscopies. The two observers independently evaluated the image quality. Lesions with insufficient images for scoring the repair tissue quality were excluded. Additionally, disagreements were re-evaluated and graded according to consensus. Patient identification details and time of follow-up were concealed for blinded assessment of the repair sites.

Altogether 104 individual observations of repaired cartilage lesions were included for the assessment of repair tissue quality. These observations were acquired from 93 separate arthroscopies from a total number of 72 lesions (Table 2). One lesion was analyzed from 52 patients and two lesions from 10 patients. The same lesion was documented on average 1.6 (range 1-5) times in different arthroscopies. Of the 62 patients, 33 were males and 29 females. The mean age of the patients was 32.1 years (SD 9.3) and the mean BMI was 25.5 kg/m^2 (SD 3.6).

The study was approved by the ethics committee of Jyväskylä Central Hospital and all patients provided informed consent at the time of index surgery.

Description of surgery

ACI was performed as described by Brittberg et al.[21] Briefly, a cartilage biopsy was first obtained during knee arthroscopy. Chondrocytes were isolated and cultured to create an autologous chondrocyte suspension in a cell culture laboratory (Sahlgrenska University Hospital, Gothenburg, Sweden). An open knee surgery was performed to repair the cartilage lesion. The lesion was debrided to subchondral bone. A periosteum patch was sutured to the surrounding cartilage margins. The seam was then finished with fibrin glue (TISSEEL, Baxter, Deerfield, Illinois, USA) to create a watertight cover over the lesion. The chondrocyte suspension was injected under the periosteum and the seam was closed with a final suture and fibrin glue.

Statistics

Internal consistency was estimated by calculating Cronbach's alpha internal consistency with bias-corrected bootstrap 95% confidence intervals (CI). The inter-rater reliability was determined by the weighted Kappa statistic (κ), intraclass correlation coefficients

(ICCs) using the one-way random effects model, and the Bland-Altman plotting method. The relative difference between observers and the absolute difference of individual items were modeled using generalized linear models with appropriate distribution and link function. The relative difference between observers was analyzed by plotting the average of the difference between the observers divided by the maximum points of the respective score with the lesion size. The level of agreement (κ) was considered as poor (κ<0.20), fair (κ=0.21-0.40), moderate (κ=0.41-0.60), substantial (κ=0.61-0.80), or very good (κ>0.8).[22] Correlation coefficients with 95% CI were calculated using Spearman's correlation with Sidak adjusted probabilities. All statistical analyses were performed with Stata version 15.0 (StataCorp, College Station, TX).

Results

The mean ICRS score was 8.8 (SD 3.6) and the mean OAS score was 5.8 (SD 2.8) for all observations. A correlation was observed between the total ICRS and OAS scores and between the individual items of the scores (Figure 1 and Table 3). The Spearman correlation curve between the total ICRS and OAS scores was nearly linear (Figure 1). The correlation was strongest between items measuring similar attributes of repair tissue. The mean values of the items were comparable to each other within each score (Table 3).

The internal consistency (Cronbach's alpha) was 0.88 (95% CI: 0.80 to 0.92) for the ICRS score and 0.79 (95% CI: 0.70 to 0.86) for the OAS score. The inter-rater reliability according to intraclass correlation coefficient was 0.89 for the ICRS score and 0.81 for the OAS score. The inter-rater reliability according to Cohen's κ for each individual item of the ICRS score ranged from 0.73 to 0.85; the range for the items of the OAS score was from 0.52 to 0.77 (Table 4).

Inter-rater reliability of both ICRS and OAS scores depended on repair size and lesion quality. This relationship was similar in both scores. The magnitude of disagreement between the observers increased according to repair size (Figure 2). In the 2 cm^2-sized lesions, the relative difference between observers was 0.04 for the ICRS score and 0.06 for the OAS score. As the repair size increased, the relative difference between observers increased up to 0.10 for the ICRS score and 0.13 for the OAS score.

The repair size affected the inter-rater reliability of the individual items. The disagreements were more frequent in larger repair sites. The highest frequency (nearly 95%) of agreement between observers was observed in small repairs. The frequency of agreement decreased to less than 80% depending on the repair size. Different items of the ICRS and the OAS scores showed variable relationships between the frequency of agreement and repair sizes. The relationship was similar between individual items of the ICRS and the OAS scores measuring the same properties of the repair tissue (Figure 3A and 3B). The observers agreed well on very low-quality and high-quality scores. The interrater agreement was lower when grading lesions with a quality between the very low and very high ends of the scales of both scores (Figure 4). Figure 5 shows representative arthroscopic images of large and small lesions with heterogeneous and homogeneous repair tissue.

Discussion

The novel finding of our study was that the inter-rater reliability of the ICRS and the OAS scores depends on the lesion size. The amount of disagreement doubled when the repair area increased from the smallest (area approximately 2 cm^2) to the largest repair sites (area >8 cm^2). All the individual items of both scores were affected by increasing lesion size. This was demonstrated by the increase in frequency of disagreement and also

by the increase of relative error between the observers. Classification of fill, integration, and appearance of repair tissue thus become ambiguous as lesion size increases. This suggests that these properties become more irregular and the repair tissue is more heterogenous in larger lesions. Both scores were able to reliably detect clearly successful and failed repairs regardless of lesion size. This information is important, as these scores are commonly used to evaluate the outcome of various cartilage repair techniques.[10-12; 23; 24] Both scores performed similarly and repair site properties affected the inter-rater reliability of both scores and their individual items equally. Our findings indicate that the reliability of both scores was well below what was expected when evaluating the repair tissue quality in large lesions.

In this study, the reliability of both scores to evaluate small lesions was relatively good. These results were consistent with previous studies. When compared with our study, previous studies evaluated score performance in more homogeneous samples. In the study of Smith et al., the validity of the ICRS and the OAS scores was evaluated using a sample of five arthroscopic video clips.[8] Van den Borne validated the use of the ICRS and OAS scores from a total of 101 arthroscopic images from arthroscopies performed 12 months after cartilage repair in patients who participated in a randomized controlled trial.[10] According to the initial report of this trial, the average and the size variation of the repair sites (2.4-2.6 cm^2; SD 1.2-1.0) were significantly smaller than in our material.[18] In an animal study, Goebel et al. analyzed macroscopically the performance of different repair tissue scoring systems, including the ICRS and the OAS scores.[20] These investigators used material of standard-sized (4 x 8 mm) lesions in 38 medial femoral condyles of Merino sheep who were sacrificed 6 months postoperatively. Due to the homogeneity of the repair sites in these studies, it was not possible to analyze the effect of altering repair site properties on the performance of the ICRS and OAS scores. Our

study differed from the other studies as ours included a wide range of lesion sizes from 104 arthroscopies.

Our study allowed a validation of score reliability in lesions of typical size for patients undergoing cartilage repair.[19] The methodological design of our study was very similar to previous trials that evaluated arthroscopic use of ICRS and OAS scores.[8; 10] According to classic reliability measures (Cronbach's alpha and ICC), our study indicated approximately similar performance of both the ICRS and OAS scores compared with previous studies performed on arthroscopic samples.[8; 10] In concordance with previous reports, the correlation between the ICRS and OAS scores was high.[8; 10; 25] As observed before, the internal consistency and the inter-rater reliability was higher for the ICRS score than for the OAS score in all used measures.[10] Our results support the previous suggestion that both the ICRS and OAS scores show a satisfactory reliability for research purposes but not for individual clinical practice.[8; 10] However, our study demonstrated that the reliability of the scores deteriorates as lesion size increases. Therefore, we suggest that if these scores are used to evaluate repair quality in typical patients undergoing cartilage repair, sufficient reliability of the scores requires at least two independent observers.

Heterogeneity of repair tissue has not been previously suggested as an explanation for the reduced reliability of the ICRS or the OAS scores. Both scores were originally designed to measure properties that experts have suggested to be important for successful cartilage repair.[1; 5; 8-12] However, neither of the scores recognizes the heterogeneity of repair tissue. Our results suggest that to improve the reliability of these scores, the scores should consider this heterogeneity. This could be achieved by providing instructions that clarify how to classify an irregularly repaired lesion. Irregularities of the repair tissue can cause

confusion both in selecting the right score item and in classifying the item properly. Defects at the margins of the lesion are easily confused as deficiencies of either fill or integration. Uneven surface or fill may lead to ambiguous classification depending on the proportions of higher and lower quality tissue. Additionally, reliability could be improved by adding new score items that recognize the heterogeneity of the repair tissue. Experts have suggested that the proportion of lesion fill is the most important measure of repair quality.[8] In both scores the fill item is supposed to represent the volume of repair tissue.[5; 8] However, as the fill item measures the thickness of the repair tissue, the item represents the volume of the repair tissue poorly if the lesion is filled with heterogeneous tissue. Even in such lesions an area of properly formed homogeneous repair tissue is typically identifiable. The area of this tissue could be expressed as a percentage of the lesion in an additional item. This item would therefore express the heterogeneity of the repair tissue. This properly formed repair tissue could be furthermore scored according to either of the scores.

The strength of our study was the heterogeneity of the lesions in our material, which resembles normal clinical material. The size and healing time of the repair sites had high variation in our material. Previous studies on the performance of the ICRS and OAS scores were performed on more standardized and homogeneous samples. The patient and lesion details were comparable to other reports of patients undergoing cartilage repair with ACI.[26; 27] To reduce possible bias, the observers scored the lesions independently and were blinded to the repair site and patient details. Therefore, our material and study design enabled an analysis of the relationship between repair site properties and score performance.

Our study has certain technical limitations. The study design was as close as possible to an authentic arthroscopic situation using retrospective material. This design is comparable with previous trials as are the limitations related to this method.[10] While the use of video clips would probably have been more authentic, only still images were available from our patients, since at the time of second-look arthroscopies only still images were used as documentation.[8] Reproducing the feeling of the repair tissue by probing from arthroscopic still images was artificial. Therefore, no conclusions should be made from the results that consider probing.

In conclusion, the reliability of the ICRS and OAS scores deteriorated as lesion size increased. Both scores performed similarly and were affected equally by lesion properties. To improve the reliability of these scores the heterogeneity of the repair tissue should be considered in the score design and instructions for use. We recommend using two observers for arthroscopic grading of repair tissue quality with these scores.

Acknowledgements

We thank Saara-Maija Hinkkanen for organizing and gathering the arthroscopic images and Jani Saalamo for digitizing the arthroscopic images for research purposes.

Role of funding source

The study was supported by the Department of Orthopaedics and Traumatology of the Helsinki University Hospital.

Competing interests

The authors declare no competing interests.

References

1. Dzioba RB. 1988. The Classification and Treatment of Acute Articular Cartilage Lesions. Arthroscopy 4:9.

2. Ahsan T, Sah RL. 1999. Biomechanics of integrative cartilage repair. Osteoarthritis Cartilage 7:29-40.

3. Riyami M, Rolf C. 2009. Evaluation of microfracture of traumatic chondral injuries to the knee in professional football and rugby players. J Orthop Surg Res 4:13.

4. Brun P, Dickinson SC, Zavan B, et al. 2008. Characteristics of repair tissue in second-look and third-look biopsies from patients treated with engineered cartilage- relationship to symptomatology and time after implantation. Arthritis Res Ther 10.

5. Brittberg M, Winalski CS. 2003. Evaluation of cartilage injuries and repair. J Bone Joint Surg Am 85-A Suppl 2:58-69.

6. Oakley SP, Portek I, Szomor Z, et al. 2005. Arthroscopy -- a potential "gold standard" for the diagnosis of the chondropathy of early osteoarthritis. Osteoarthritis Cartilage 13:368-378.

7. Oakley SP, Lassere MN. 2003. A critical appraisal of quantitative arthroscopy as an outcome measure in osteoarthritis of the knee. Semin Arthritis Rheum 33:83-105.

8. Smith GD, Taylor J, Almqvist KF, et al. 2005. Arthroscopic assessment of cartilage repair: a validation study of 2 scoring systems. Arthroscopy 21:1462-1467.

9. Peterson L, Minas T, Brittberg M, et al. 2000. Two- to 9-year outcome after autologous chondrocyte transplantation of the knee. Clin Orthop Relat Res (374):212-234.

10. van den Borne MP, Raijmakers NJ, Vanlauwe J, et al. 2007. International Cartilage Repair Society (ICRS) and Oswestry macroscopic cartilage evaluation scores validated for use in Autologous Chondrocyte Implantation (ACI) and microfracture. Osteoarthritis Cartilage 15:1397-1402.

11. Knutsen G, Engebretsen L, Ludvigsen TC, et al. 2004. Autologous chondrocyte implantation compared with microfracture in the knee. A randomized trial. The Journal of bone and joint surgeryAmerican volume 86-A:455-464.

12. Henderson I, Francisco R, Oakes B, et al. 2005. Autologous chondrocyte implantation for treatment of focal chondral defects of the knee--a clinical, arthroscopic, MRI and histologic evaluation at 2 years. The Knee 12:209-216.

13. Wilkinson L, Inference. TTFoS. 1999. Statistical methods in psychology journals - Guidelines and explanations. Am Psychol 54:11.

14. Streiner DL. 2003. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess 80:99-103.

15. Thompson B, Vacha-Haase T. 2000. Psychometrics is Datametrics: the Test is not Reliable. Educ Psychol Meas 60:22.

16. Vacha-Haase T. 1998. Reliability Generalization: Exploring Variance in Measurement Error Affecting Score Reliability Across Studies. Educ Psychol Meas 58:15.

17. Caruso JC. 2000. Reliability Generalization of the Nea Personality Scales. Educ Psychol Meas 60:19.

18. Saris DB, Vanlauwe J, Victor J, et al. 2008. Characterized chondrocyte implantation results in better structural repair when treating symptomatic cartilage defects of the knee in a randomized controlled trial versus microfracture. The American Journal of Sports Medicine 36:235-246.

19. Engen CN, Engebretsen L, Aroen A. 2010. Knee Cartilage Defect Patients Enrolled in Randomized Controlled Trials Are Not Representative of Patients in Orthopedic Practice. Cartilage 1:312-319.

20. Goebel L, Orth P, Cucchiarini M, et al. 2017. Macroscopic cartilage repair scoring of defect fill, integration and total points correlate with corresponding items in histological scoring systems - a study in adult sheep. Osteoarthritis Cartilage 25:581-588.

21. Brittberg M, Lindahl A, Nilsson A, et al. 1994. Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. N Engl J Med 331:889-895.

22. Landis JR, Koch GC. 1977. The measurement of observer agreement for categorical data. Biometrics 33:16.

23. Peterson L, Brittberg M, Kiviranta I, et al. 2002. Autologous chondrocyte transplantation. Biomechanics and long-term durability. The American Journal of Sports Medicine 30:2-12.

24. Vasara AI, Nieminen MT, Jurvelin JS, et al. 2005. Indentation stiffness of repair tissue after autologous chondrocyte transplantation. Clin Orthop Relat Res (433):233-242.

25. Goebel L, Orth P, Muller A, et al. 2012. Experimental scoring systems for macroscopic articular cartilage repair correlate with the MOCART score assessed by a high-field MRI at 9.4 T--comparative evaluation of five macroscopic scoring systems in a large animal cartilage defect model. Osteoarthritis Cartilage 20:1046-1055.

26. Peterson L, Vasiliadis HS, Brittberg M, et al. 2010. Autologous chondrocyte implantation: a long-term follow-up. The American Journal of Sports Medicine 38:1117-1124.

27. Moseley JB, Jr., Anderson AF, Browne JE, et al. 2010. Long-term durability of autologous chondrocyte implantation: a multicenter, observational study in US patients. The American Journal of Sports Medicine 38:238-246.

Figures

Figure 1. Scatter plot of correlation between ICRS and OAS scores. Spearman correlation with 95% confidence interval (CI). The size of the circle reflects the number of lesions.

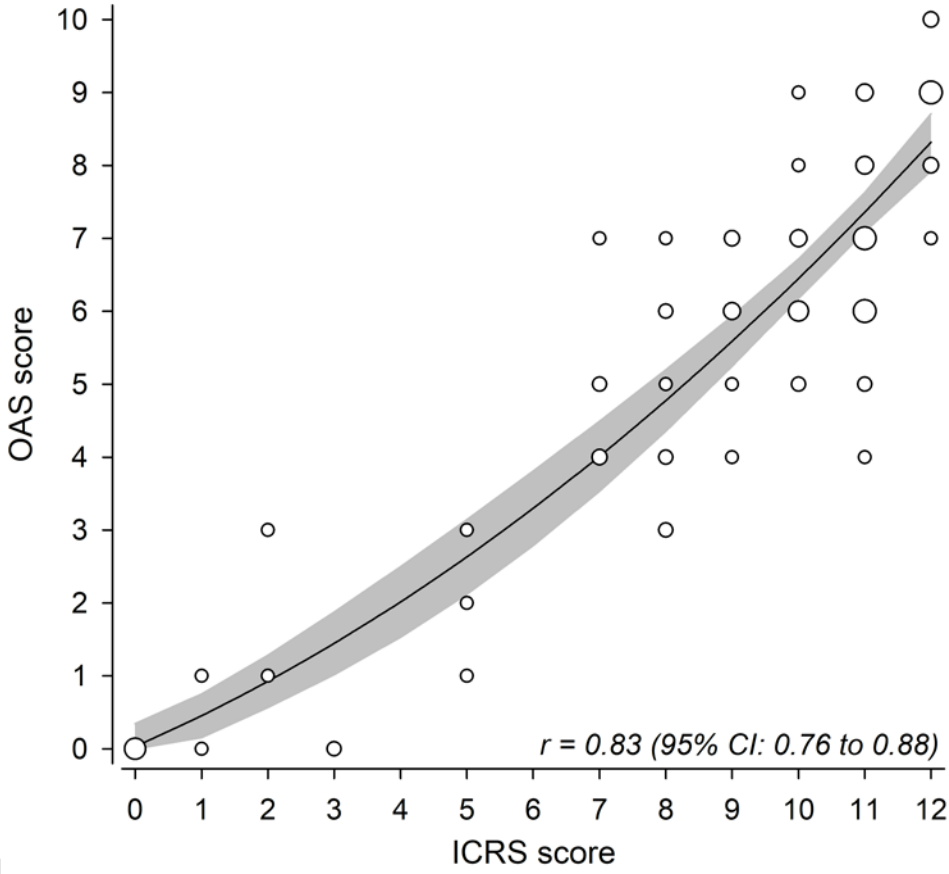

Figure 2. The relative difference between observers of the ICRS and OAS scores in relation to lesion size. The relative difference describes how large the disagreement between observers was in proportion to the maximum points of the corresponding score.

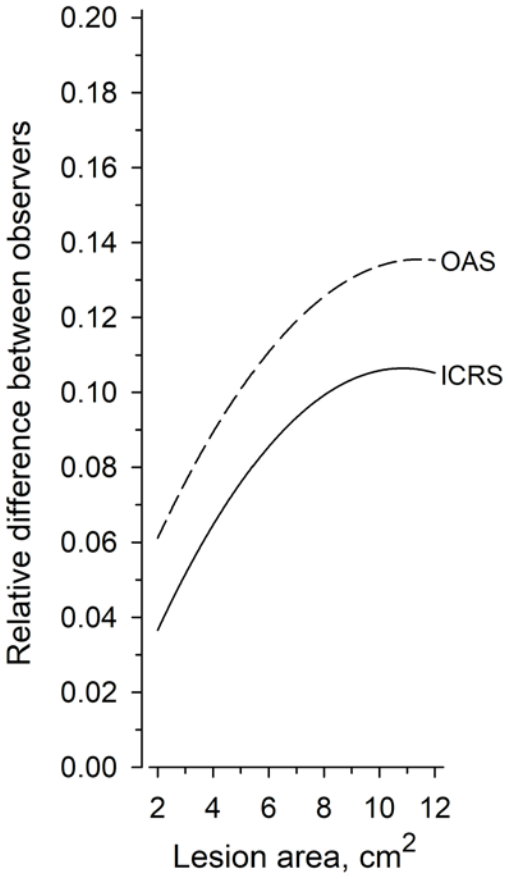

Figure 3. The absolute percentage of agreement between observers of individual items of the ICRS score (panel A) and the OAS score (panel B). Shaded area shows the 95% confidence interval.

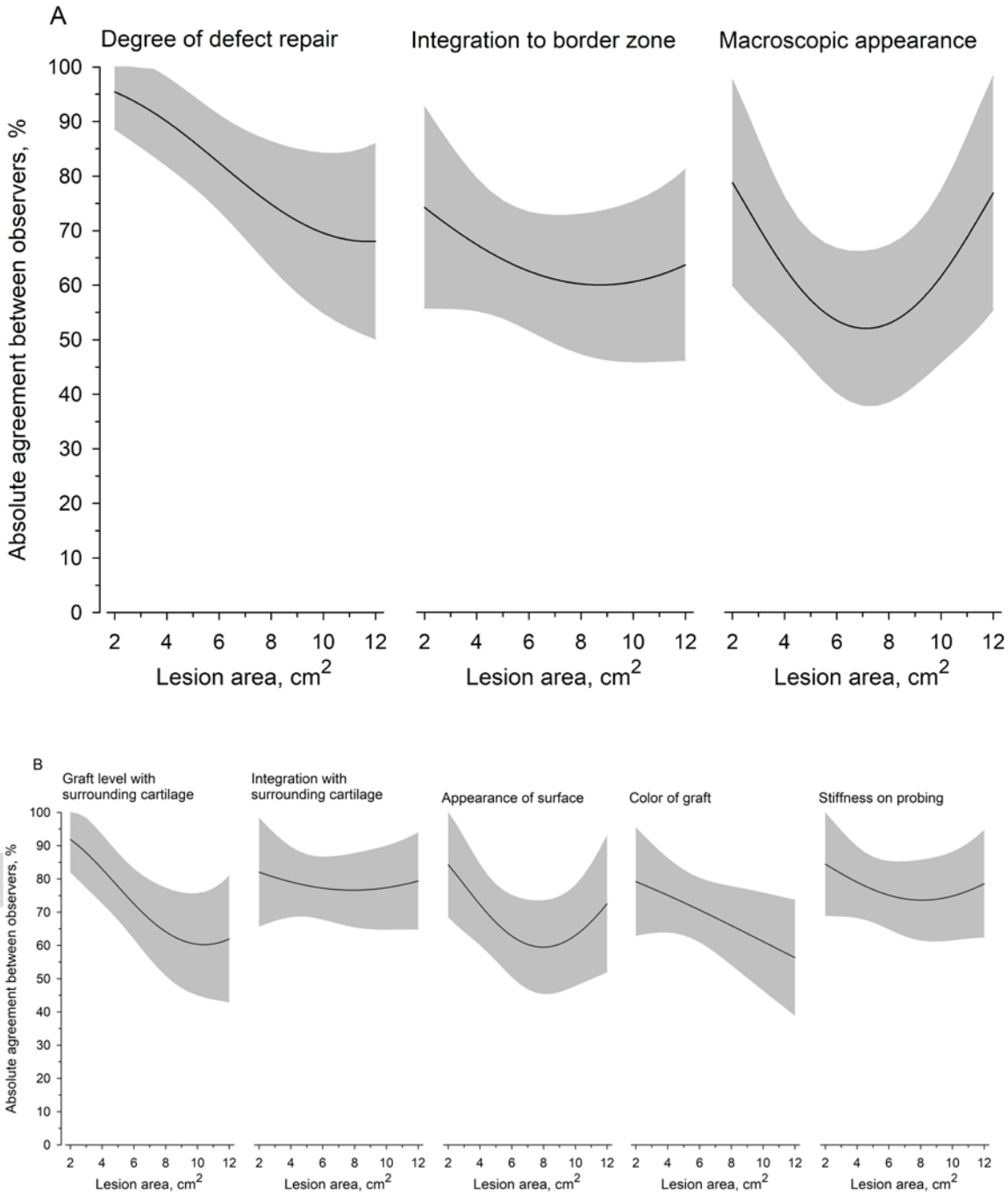

Figure 4. Bland-Altman plot comparing the scoring of two independent observers for both the ICRS and OAS scores. The dotted lines show the 95% limits of agreement. Circle size represents the number of samples.

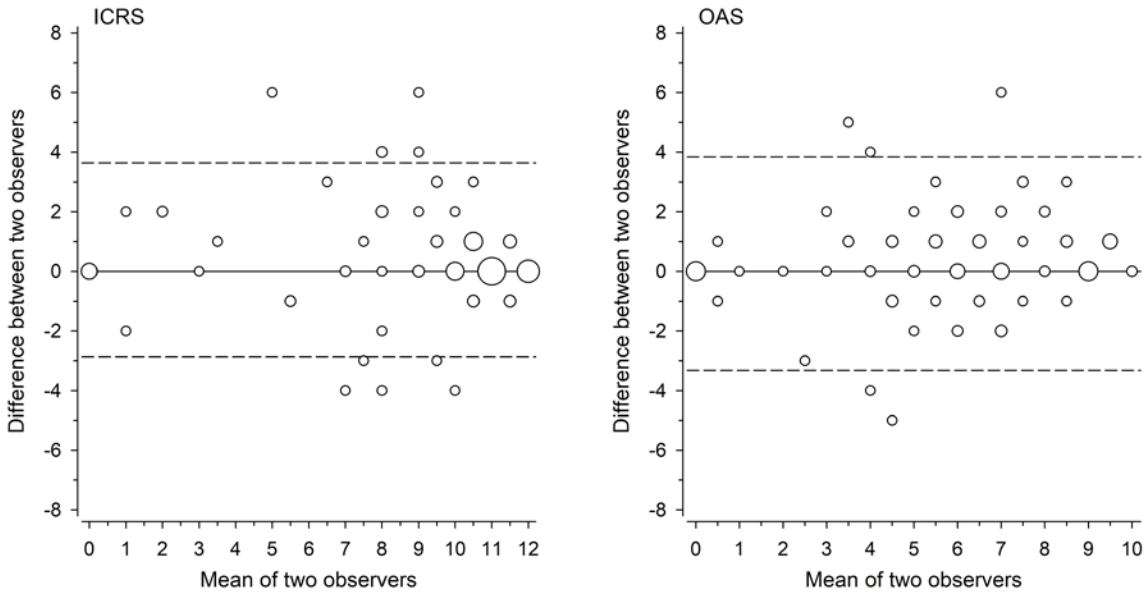

Figure 5. Second-look arthroscopic images that represent either heterogeneous or homogeneous repair tissue from four patients. (a) The repair tissue varies in thickness in a large lesion (14.3cm^2). (b) Insufficient formation of repair tissue at the margin of a large lesion (9.0cm^2). (c) Homogeneous repair tissue with fibrillated surface in a small lesion (3.0cm^2). (d) High-quality and homogeneous repair in a large lesion (9.4cm^2).

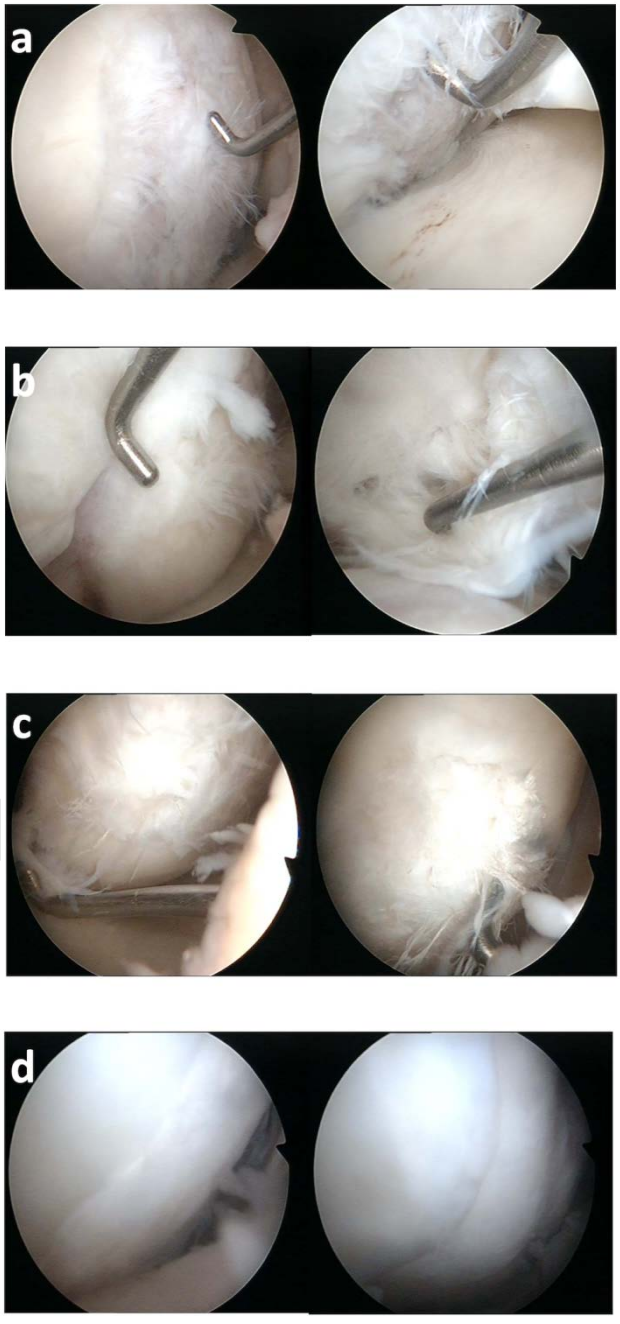

Table 1. International Cartilage Repair Society (ICRS) score (protocol A) and Oswestry Arthroscopic Score (OAS) for arthroscopic repair tissue quality assessment.[5; 8]

Score / *Item* / Scale	**Points**
ICRS (protocol A) scoring	**0-12**
Degree of defect repair (protocol A)	
Level with surrounding cartilage	4
75% repair of defect depth	3
50% repair of defect depth	2
25% repair of defect depth	1
0% repair of defect depth	0
Integration to border zone	
Complete integration with surrounding cartilage	4
Demarcating border <1 mm	3
2/4 of graft integrated, 1/4 with a notable border >1 mm width	2
1/2 of graft integrated with surrounding cartilage, 1/2 with a notable border >1 mm	1
From no contact to 1/4 of graft integrated with surrounding cartilage	0
Macroscopic appearance	
Intact smooth surface	4
Fibrillated surface	3
Small, scattered fissures or cracks	2
Several small or few but large fissures	1
Total degeneration of grafted area	0
OAS scoring	**0-10**
Graft level with surrounding cartilage	
Level	2
Raised	1
Below	0
Integration with surrounding cartilage	
Complete	2
Minor disruption (<25% of area)	1
Major disruption (>25% of area)	0
Appearance of surface	
Smooth	2
Fine fronds	1
Severe fronds/fibrillation	0
Color of graft	
Pearly, hyaline-like	2
White	1
Yellow bone	0
Stiffness on probing	
Normal compared with adjacent cartilage	2
Softer	1
Very soft/hard	0

Table 2. Lesion characteristics

	Measures
Etiology, n (%):	
Osteochondritis dissecans (OCD)	17 (24)
Trauma	34 (47)
Non-trauma	21 (29)
Lesion location, n (%):	
Medial femoral condyle	37 (51)
Lateral femoral condyle	10 (14)
Trochlea of femur	15 (21)
Lateral tibial condyle	1 (1)
Patella	9 (13)
Lesion size cm^2, mean (range)	6.9 (1.6-21.8)
Lesion size, cm^2, n (%):	
<4	19 (26)
4-6	16 (22)
6-8	12 (17)
8-10	12 (17)
>10	13 (18)

Table 3. Spearman correlation coefficients between the individual items of the ICRS and the OAS scores.

Score	**ICRS**		
	Degree of defect repair	Integration to border zone	Macroscopic appearance

		r (95% CI)	r (95% CI)	r (95% CI)
OAS	Graft level with surrounding cartilage	0.69 (0.58-0.78)***	0.25 (0.06-0.43)	0.53 (0.38-0.66)***
	Integration with surrounding cartilage	0.47 (0.31-0.61)***	0.91 (0.88-0.94) ***	0.38 (0.20-0.54)***
	Appearance of surface	0.57 (0.42-0.68)***	0.36 (0.18-0.52)**	0.92 (0.88-0.94)***
	Color of graft	0.51 (0.36-0.64)***	0.45 (0.29-0.59)***	0.58 (0.44-0.70)***
	Stiffness in probing	0.51 (0.35-0.64)***	0.36 (0.18-0.52)**	0.59 (0.44-0.70)***

*p<0.05, **p<0.01, ***p<0.001; Sidak-adjusted probabilities.

ICRS, International Cartilage Repair Society; OAS, Oswestry Arthroscopic Score

Table 4. Inter-rater reliability for the total scores and individual items of the ICRS and OAS scores.

Score / Item	Mean (SD)	κ (95% CI)	ICC (95% CI)
ICRS score	8.8 (3.6)		0.89 (0.84-0.92)
Degree of defect repair	3.3 (1.3)	0.85 (0.75-0.93)	
Integration to border zone	2.8 (1.4)	0.75 (0.62-0.86)	
Macroscopic appearance	2.8 (1.3)	0.73 (0.57-0.83)	
OAS score	5.8 (2.8)		0.81 (0.74-0.87)
Graft level with surrounding cartilage	1.2 (0.8)	0.60 (0.47-0.73)	
Integration with surrounding cartilage	1.2 (0.8)	0.68 (0.53-0.81)	
Appearance of surface	1.1 (0.8)	0.67 (0.53-0.77)	
Color of graft	1.0 (0.5)	0.52 (0.36-0.67)	
Stiffness on probing	1.2 (0.8)	0.77 (0.66-0.85)	

ICRS, International Cartilage Repair Society; OAS, Oswestry Arthroscopic Score